Artificial and Computational Intelligence

AIMLCZG557

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

Presented by
V Indumathi –Guest Faculty –BITS-WILP
indumathi.p@wilp.bits-pilani.ac.in

AIMLCZG557 – CS#2

M1: Introduction to AI & Intelligent Agents & Environments

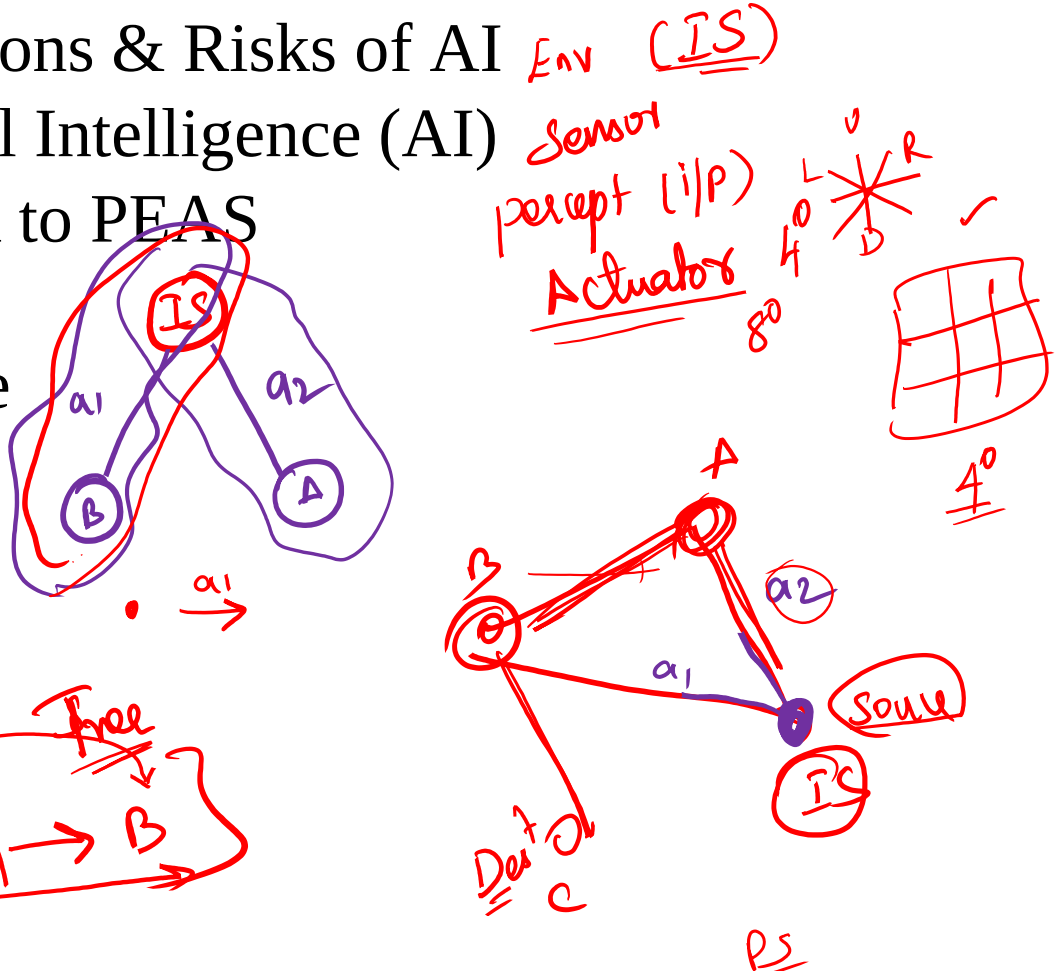
ACI: Disclaimer and Acknowledgements

⚠

- Few content for these slides may have been obtained from prescribed books and various other source(s) on the Internet.
- We hereby acknowledge all the contributors for their material and inputs and gratefully acknowledge people others who made their course materials freely available online. We have provided source information wherever necessary.
- We have added and modified the content to suit the requirements of the class dynamics & live session's lecture delivery flow for presentation.
- This *is not a full-fledged* reading materials. Students are requested to refer to the textbook w.r.t detailed content as per the course handout as shared over e-learning portal - Taxila.
- **Slide Source / Preparation / Review:**
- From BITS Pilani WILP**:** Prof. Rajavadhana, Prof. Indumathi, Prof. Sangeetha and Prof Parthasarathy PD
- From BITS On-campus & External : Mr.Santosh GSK

Agenda for CS #2

1) Recap of CS#1
3) Foundations & applications & Risks of AI
4) Perspectives to Artificial Intelligence (AI)
5) Agents and Introduction to PEAS
6) PEAS
   - Performance measure
   - Environment
   - Actuators
   - Sensors
7) Environment types
8) Agent types Intro
9) Q&A!



AI Application Areas

- Healthcare
- Finance
- Education
- Automotive
- Retail
- Manufacturing
- Customer Service
- Natural Language Processing (NLP)
- Computer Vision
- Robotics
- Smart Cities
- Agriculture
- Entertainment
- Cybersecurity
- Environmental Conservation
- Culinary
- Transportation
- …
- ……

Can you give one example in each area how AI is used?

Areas Contributing to AI

Philosophy
Mathematics
Economics
Neuroscience
Psychology
Computer Engineering
Control theory, Cybernetics
Linguistics

Glance Textbook, T1: Section 1.2 to know more!

Foundations of AI

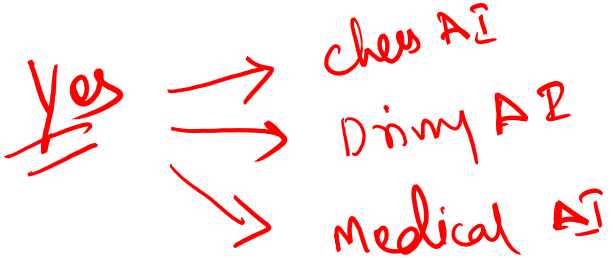
- Philosophy
  - Logic, method of reasoning
  - Mind as a physical state that operates as a set of rules
  - Foundations of learning, language, rationality
- Mathematics
  - Logic (propositional, first order etc.), formula representation and proof
  - Computation, algorithms
- Economics
  - Formal theory of rational decisions
  - Combined decision theory and probability theory for decision making under uncertainty
  - Game Theory and Markov Decision process
- Neuroscience
  - Study of brain functioning
  - How brains compare to computers
- Psychology
  - How do we think and act?
  - Cognitive psychology – perceives the brain as an information processing machine
  - Cognitive science – How can computer models be used to study language, memory and thinking from a psychology perspective
- Computer Science
  - Delves into how to build powerful machines to make AI possible
  - Rapid advancement in computational processing power

What is intelligence?

- Dictionary.com: *capacity for learning, reasoning, understanding, and similar forms of mental activity.*
- Ability to perceive and act in the world
- *Reasoning*: proving theorems, medical diagnosis
- *Planning*: take decisions
- *Learning and Adaptation*: recommend movies, learn traffic patterns
- *Understanding*: text, speech, visual scene

Intelligence vs. humans

- Are humans intelligent?
  - replicating human behavior early hallmark of intelligence
- Are humans always intelligent?
- Can non-human behavior be intelligent?



What is *artificial* intelligence?

human-like *vs.* rational

thought *vs.* behaviour

Thinking Humanly "The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	**Thinking Rationally** "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** "Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

What is *artificial* intelligence?

human-like *vs.* rational

thought *vs.* behaviour		
	Systems that think like humans	Systems that think rationally
	Systems that act like humans	Systems that act rationally

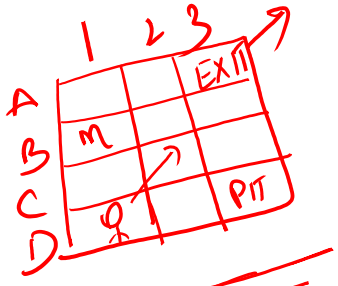
Thought/ Reasoning	**Human Performance**	**Rational Performance**
	(1) Thinking Humanly	(3) Thinking Rationally
Acting	(2) Acting Humanly	(4) Acting Rationally

What is *artificial* intelligence?

	Human Performance	Rational Performance
Thought/ Reasoning	(1) Thinking Humanly	(3) Thinking Rationally
Acting	(2) Acting Humanly	(4) Acting Rationally

T1fn(D1, UP) → C1

T2fn (D, R) → D2



IS : D1

GS : A3

f0 Action: MUP, D, Lef, R1

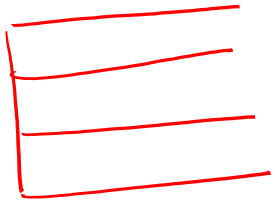
Obs - B1 & D3

Tree

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

What is *artificial* intelligence?

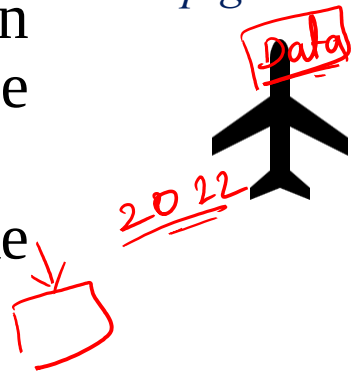
	Human Performance	Rational Performance
Thought/ Reasoning	(1) Thinking Humanly	(3) Thinking Rationally
Acting	(2) Acting Humanly	(4) Acting Rationally



(1) Thinking Humanly : Cognitive Modelling Approach

- Uses Cognitive Science
  - Very hard to understand how humans think.
  - Usually we use post-facto rationalizations, irrationality of human thinking.
- How do we capture human thinking ?
  - Introspection
  - Psychological experiment's
  - Brain imaging
- Do we want a machine that beats humans in chess or a machine that *thinks like humans* while beating humans in chess?
  - Deep Blue supposedly *DOESN'T* think like humans..

Aeronautical engineering texts do not define the goal of their field as making "machines that fly so exactly like pigeons that they can fool even other pigeons!"



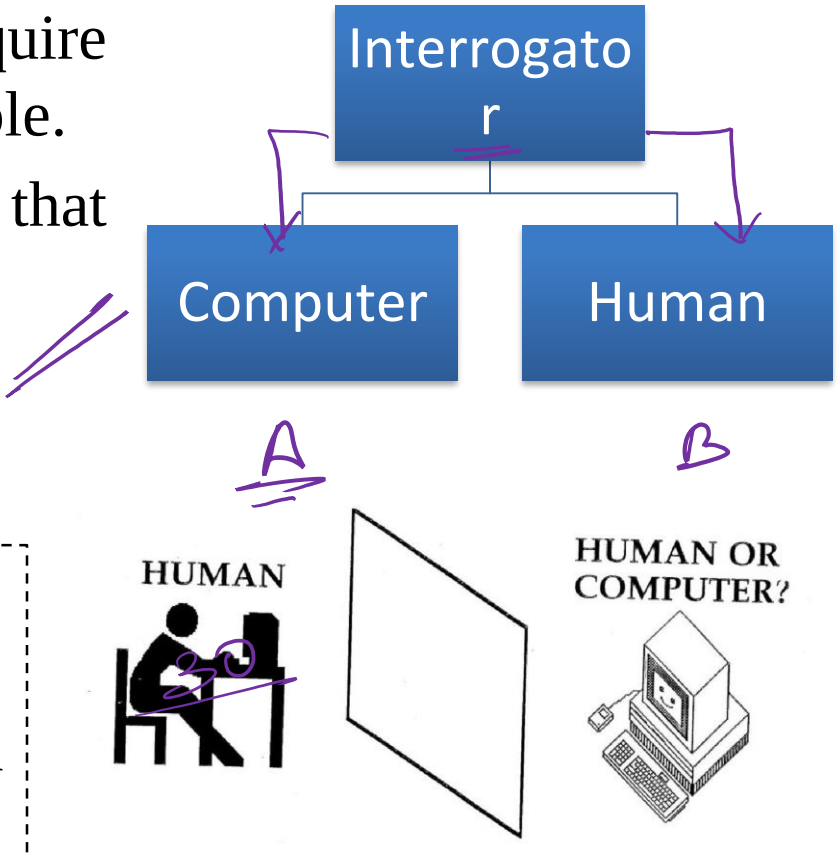
(1) Thinking Humanly : Cognitive Modelling Approach

innovate achieve lead

- Thinking like humans important in Cognitive Science applications
  - Intelligent tutoring
  - Expressing emotions in interfaces… HCI
- Growth of Cognitive science and AI supports each other.
- "General Problem Solver" (Newell and Simon, 1961)
  - Designed to work as a universal problem solver
  - Problems represented by horn clauses
  - First AI Machine which has KB + Inference separation
  - Authors focus on this is on comparing the trace of its reasoning steps to traces of human subjects solving the same problems

(2) Acting Humanly : Turing Test Approach

- It is about creating machines that can act or perform tasks which would require intelligence when performed by people.
- Assume you have built a machine that can Act Humanly.
- Now, let us test it!
- Turing Test



If the human/interrogator cannot tell whether the responses from the other side of a wall are coming from a human or computer, then the computer is intelligent.

(2) Acting Humanly : Turing Test Approach

- Skills necessary to pass these tests:
  - NLP, Knowledge Representation, Automated Reasoning, ML + Computer Vision & Robotics (for total turing test).
- Loebner Prize
  - Every year in Boston
  - Expertise-dependent tests: limited conversation
- Chabot - Eugene Goostman - Pretended to be a thirteen-year-old Ukrainian boy:
  - Passed the turing test for the first time! 10/30 Judges believed the response is from human
- What if people call a human, a machine?
  - Shakespeare expert / Shakuntala Devi!
  - Make human-like errors
- Problems
  - Not reproducible, constructive or mathematically analyzable

(3) Thinking Rationally : Laws of Thought Approach

- *Aristotle*: what are correct arguments/thought processes? Logic!
- Invention of Formal Logic, Greek Philosopher Aristotle, Third century BC.
- Introduced syllogisms, providing argument structures..

In all boring classes, students sleep

It is a boring class

Students sleep in this class [Are you ?]

- Field of Logics gave rise to *codifying* rational thinking
  - When elements are 'things', we reason about things
- Problems? (1) Not all intelligent behavior is mediated by logical deliberation (reflexes) (2) It is computationally exhaustive to build systems even with a few hundred facts. Hence, such systems are not scalable.

(4) Acting Rationally: The Rational Agent Approach

- Acting Rationally is about building *Rational Agents*.
- An agent is an entity that perceives and acts! (This course is majorly about designing rational agents).
- Agent – Something which acts, i.e., performs an action on the environment.
- *Rational Agent* - One that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome
- Computational limitations make perfect rationality unachievable.
- Design best program for given machine resources

(4) Acting Rationally: The Rational Agent Approach

innovate achieve lead

- Rational behaviour: doing the *right thing*
- The *right thing*: that which is expected to maximize goal achievement, given the available information
- Rationality can include
  - Correct inference – using logical reasoning to perform action in order to achieve a goal, e.g., when to brake a car
  - Make an inference even when there is no correct inference, e.g., predicting an action in self-driving car during an accident to minimize life damage
  - Acting without inferring (need not always be deliberative), e.g., recoiling from a hot stove is a reflex action without inference
- Rational behaviour is not just about correct inference / thinking, skills needed to pass turing test etc.

Risks of AI

- Lethal autonomous weapons: These are defined by the UN as weapons that can locate, select, and eliminate human targets *without* human intervention.
  - A primary concern with such weapons is their *scalability*: the absence of a requirement for human supervision means that a small group can deploy an arbitrarily large number of weapons against human targets defined by any feasible recognition criterion.
- Surveillance and persuasion: AI can be used in a scalable fashion to perform mass surveillance of individuals and detect activities of interest.
  - By tailoring information flows to individuals through social media, based on machine learning techniques, political behavior can be modified and controlled to some extent.

Risks of AI

- Biased decision making: Careless or deliberate misuse of machine learning algorithms for tasks such as evaluating parole and loan applications can result in decisions that are biased by race, gender, or other protected categories. Often, the data themselves reflect pervasive bias in society.
- Impact on employment: AI do some of the tasks that humans might otherwise do, but they also make humans more productive and therefore more employable, and make companies more profitable and therefore able to pay higher wages.
  - AI use generally results in increasing wealth but tends to have the effect of shifting wealth from labor to capital, further exacerbating increases in inequality.

Risks of AI

- Safety-critical applications: As AI techniques advance, they are increasingly used in high-stakes, safety-critical applications such as driving cars and managing the water supplies of cities.
  - The field of AI will need to develop technical and ethical standards at least comparable to those prevalent in other engineering and healthcare disciplines where people's lives are at stake.
- Cybersecurity: AI techniques are useful in defending against cyberattack, for example by detecting unusual patterns of behavior, but they will also contribute to the potency, survivability, and proliferation capability of malware.
  - For example, reinforcement learning methods have been used to create highly effective tools for automated, personalized blackmail and phishing attacks.
- Other risks such as sustainability, deepfakes etc…

Agents

- An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators.
- *Human agent:*
  - eyes, ears, and other organs for sensors;
  - hands, legs, mouth, and other body parts for actuators
- *Robotic agent:*
  - cameras and infrared range finders for sensors;
  - various motors for actuators
- *AI agent:*
  - AI agents are intelligent entities powered by LLMs capable of **perceiving** environments, **reasoning** about goals, and **executing** actions.
  - AI agent **senses** its environment, **processes** info, makes **decisions**, and takes **actions** aligned with achieving **assigned** objectives.

AI Agents vs Agentic AI

- Further in this course, we will treat an agent as a *rational* agent and not go into specifics if it's a robot, AI agent, multiple agents etc.
- Just for clarity, while AI agent and Agentic AI are related, there is a difference.
- It all lies in the level of autonomy!
- A customer support bot that answers FAQs using predefined scripts or learnt logic or reasoned rules is an **AI agent.**
- But an example of Agentic AI could be, A network of bots that not only answer FAQs but also escalate issues, analyze customer sentiment, adapt responses, and coordinate with backend systems to resolve problems autonomously. The bifurcation of task, orchestration & pipelining is also learnt by the agents autonomously where reasoning is predominantly using language models (natural Languages).

Examples of Agents

Physically grounded agents

- Intelligent buildings
- Autonomous spacecraft

Softbots

- Expert Systems
- IBM Watson
- ChatGPT/Google?

IBM Watson

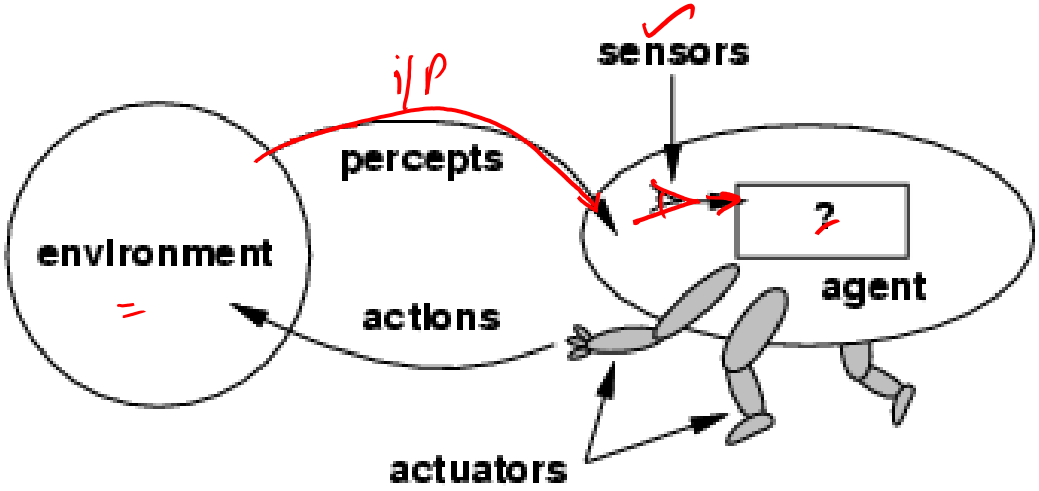
Claude

alexa

ChatGPT

Agents & Environments

- Intelligent agents have sensors, effectors.
- Implement mapping from percept sequence to actions



- The agent function maps from percept histories to actions:

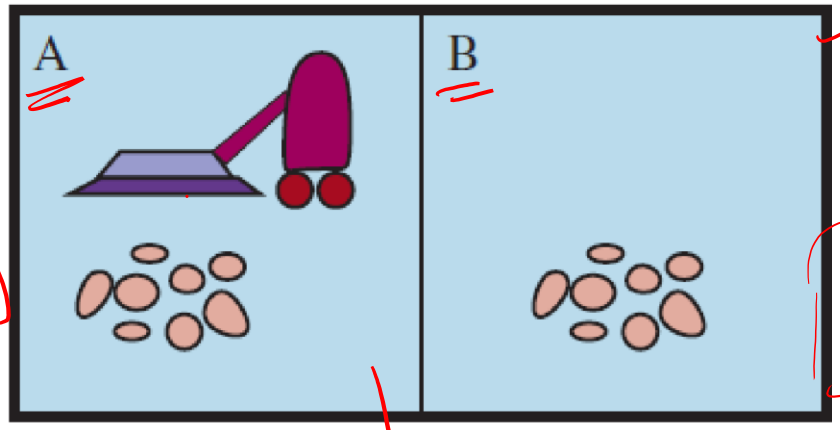
$$[f: \mathcal{P}^* \rightarrow \mathcal{A}]$$

- The agent program runs on the physical architecture to produce f

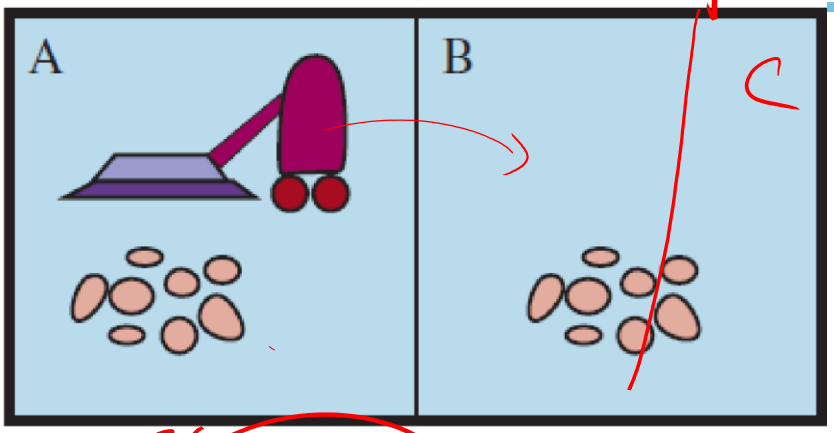
agent = architecture + program

A vacuum-cleaner agent

- Percepts: location and contents, e.g., [A , Dirty]
- Actions: *Left*, *Right*, *Suck*, *NoOp*



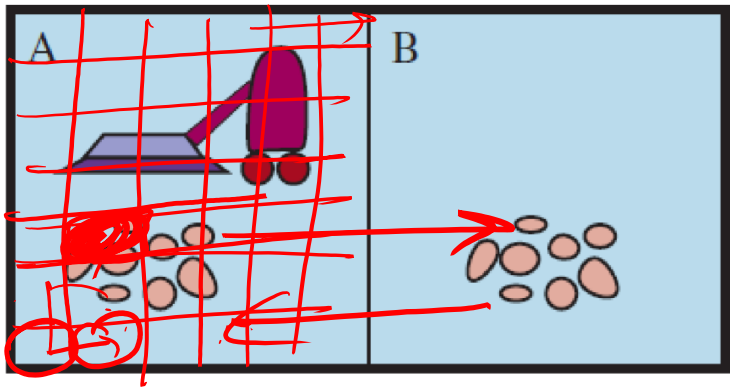
A vacuum-cleaner agent



innovate achieve lead

A vacuum-cleaner agent

Percept sequence	Action
$[A, Clean]$	$Right$
$[A, Dirty]$	$Suck$
$[B, Clean]$	$Left$
$[B, Dirty]$	$Suck$
$[A, Clean], [A, Clean]$	$Right$
$[A, Clean], [A, Dirty]$	$Suck$
⋮	⋮
$[A, Clean], [A, Clean], [A, Clean]$	$Right$
$[A, Clean], [A, Clean], [A, Dirty]$	$Suck$
⋮	⋮



Rational Agents

- An agent should strive to *do the right thing*, based on what it can perceive and the actions it can perform.
- The right action is the one that will cause the agent to be most successful!
- Performance measure: An objective criterion for success of an agent's behavior.
- E.g., performance measure of a vacuum-cleaner agent could be
  - amount of dirt cleaned up,
  - amount of time taken,
  - amount of electricity consumed,
  - amount of noise generated, etc.

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

Ideal Rational Agents

"For each possible percept sequence, does whatever action is expected to maximize its performance measure on the basis of evidence perceived so far and built-in knowledge."

- Rationality is *distinct* from **omniscience**
  - Omniscience (all-knowing with infinite knowledge)
- Agents can perform actions in order to modify future percepts so as to obtain useful information (information gathering, exploration)
- An agent is *autonomous* if its behavior is determined by its own experience (with ability to learn and adapt)

Are ideal rational agents possible?

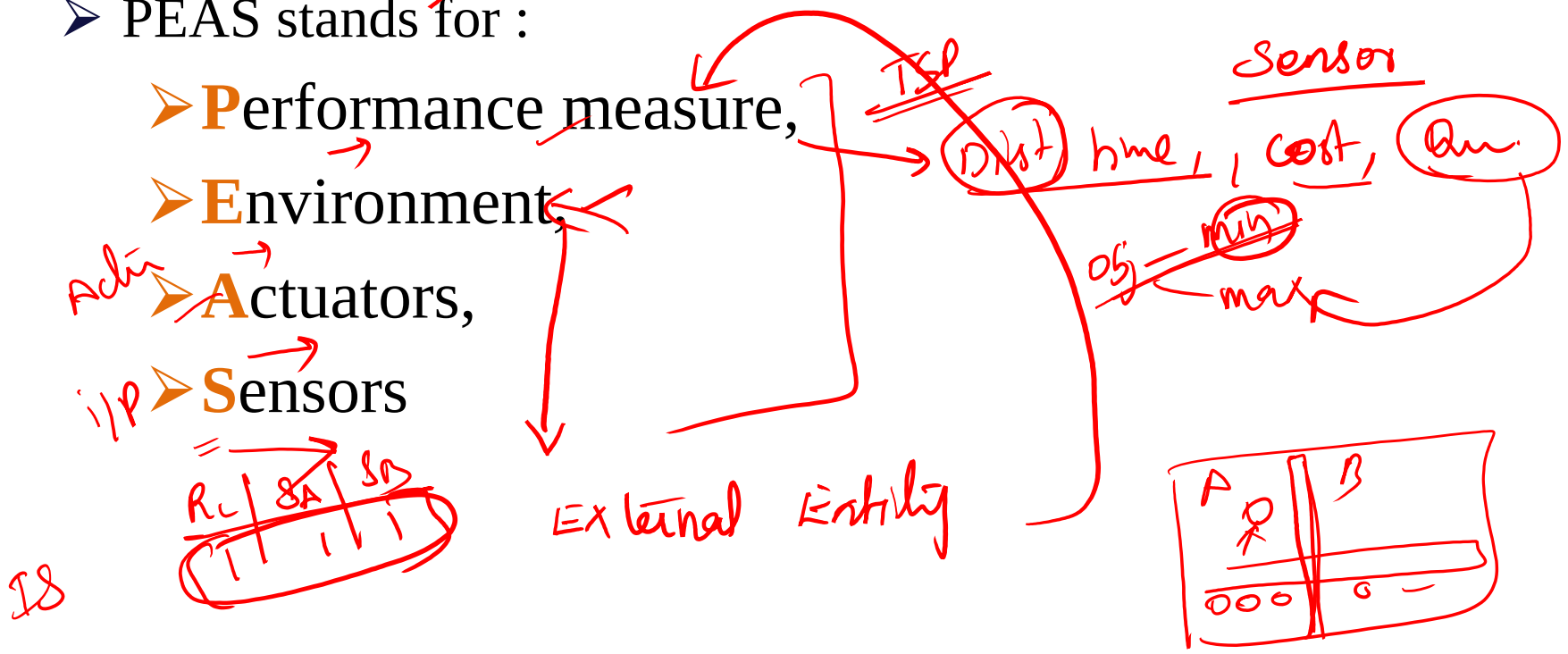
Bounded Rationality

- We have a performance measure to optimize
- Given our state of knowledge
- Choose optimal action
- *Given limited computational resources*

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

PEAS: Specifying Task Environments

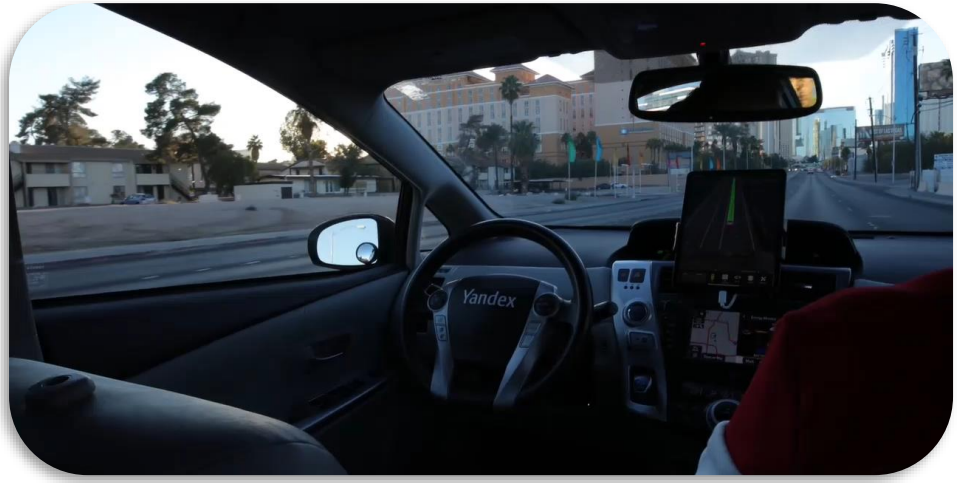
- Must first specify the setting for intelligent agent design and the PEAS technique is often used to do so.
- PEAS stands for :
  - **P**erformance measure,
  - **E**nvironment,
  - **A**ctuators,
  - **S**ensors



BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

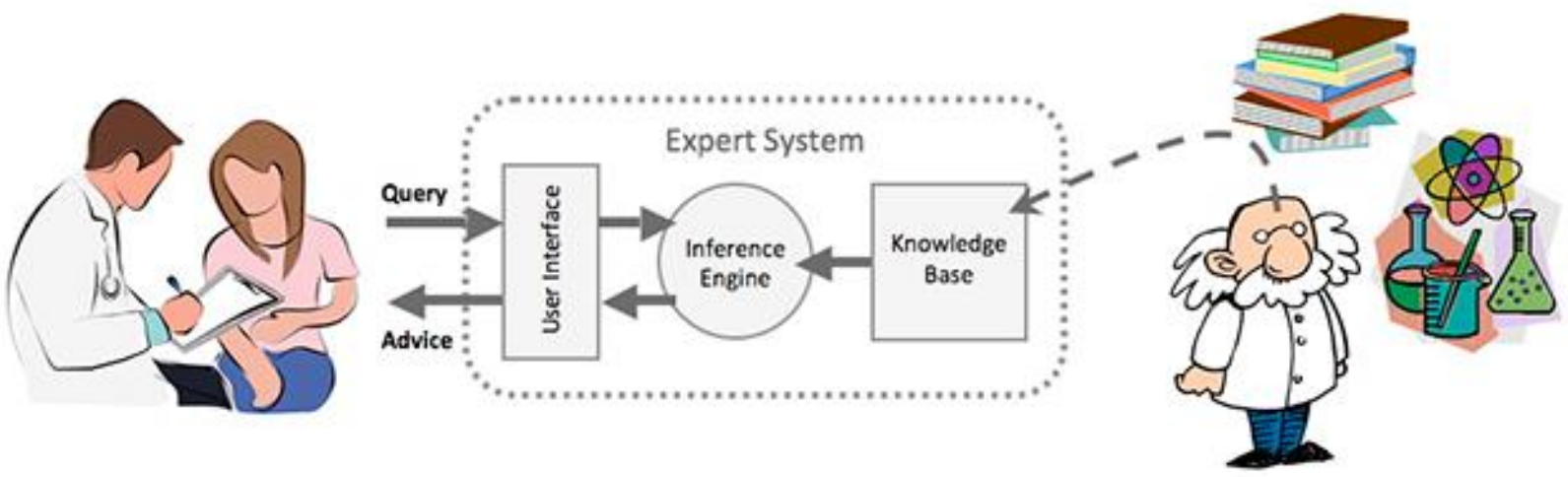
PEAS for Automated taxi driver

- *Performance measure*: Safe, fast, legal, comfortable trip, maximize profits
- *Environment:* Roads, other traffic, signals, pedestrians, carts and customers.
- *Actuators:* Steering wheel, accelerator, brake, signal, horn
- *Sensors:* Cameras, sonar, speedometer, GPS, odometer, engine sensors, keyboard



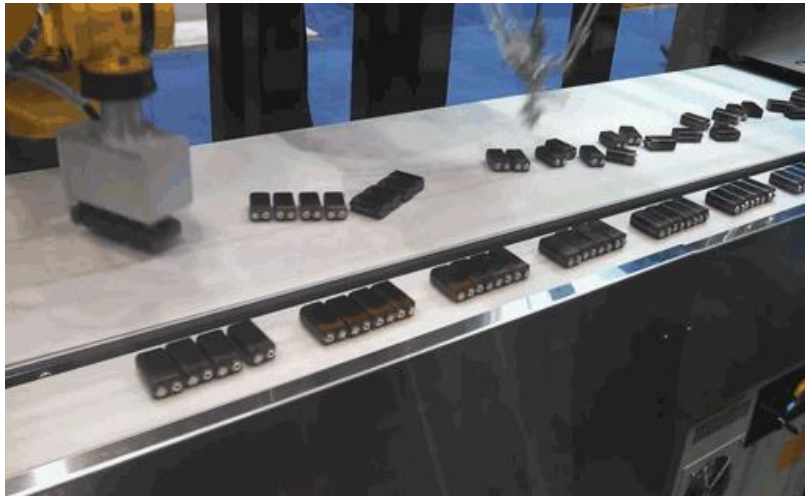
PEAS - Medical diagnosis system

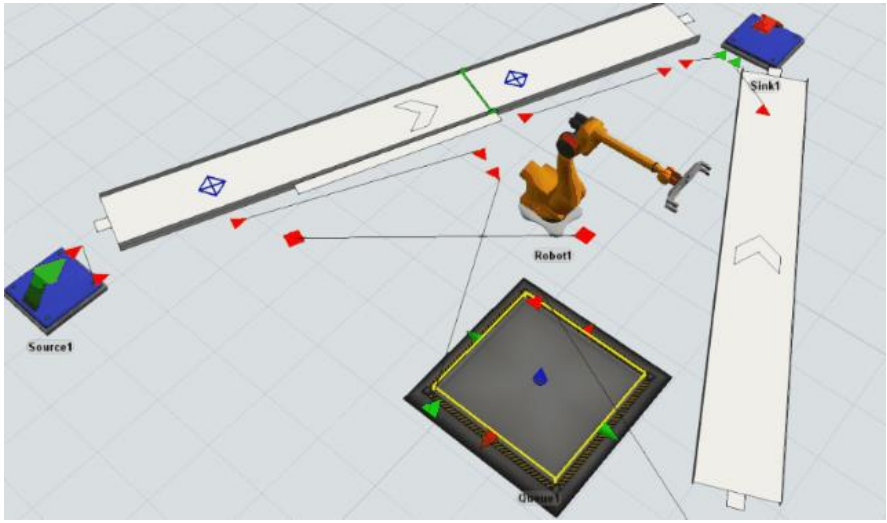
- *Performance measure*: Healthy patient, minimize costs, lawsuits
- *Environment*: Patient, hospital, staff
- *Actuators*: Screen display (questions, tests, diagnoses, treatments, referrals)
- *Sensors*: Keyboard (entry of symptoms, findings, patient's answers)



PEAS - Part-picking robot [Activity 5min]

- *Performance measure:* Percentage of parts in correct bins
- *Environment:* Conveyor belt with parts, bins
- *Actuators:* Jointed arm and hand
- *Sensors:* Camera, joint angle sensors



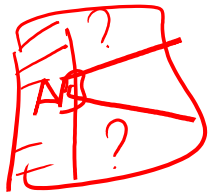


Properties of Environments

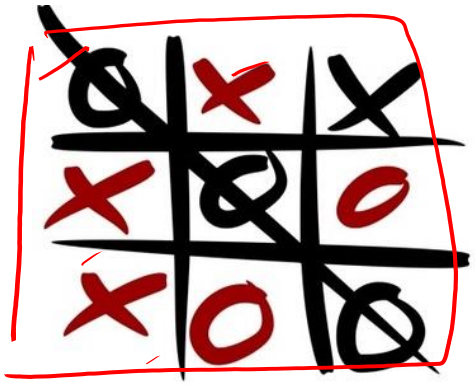
- Observability: full *vs.* partial *vs. non*
- Deterministic *vs.* stochastic
- Episodic *vs.* sequential
- Static *vs. Semi-dynamic vs.* dynamic
- Discrete *vs.* continuous
- Single Agent *vs.* Multi Agent (Cooperative, Competitive, Self-Interested)

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

Observability



- Fully observable (vs. partially observable): An agent's sensors give it access to the complete state of the environment at each point in time.

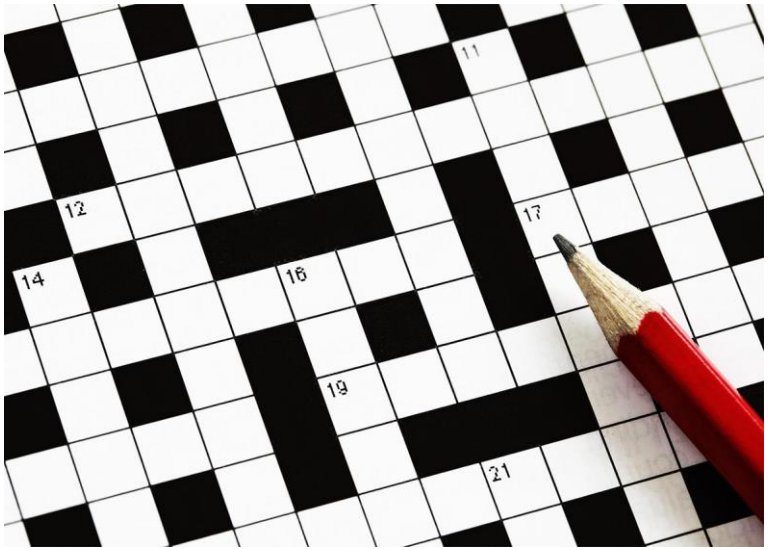


Fully Observable: Tic-tac-Toe and Chess

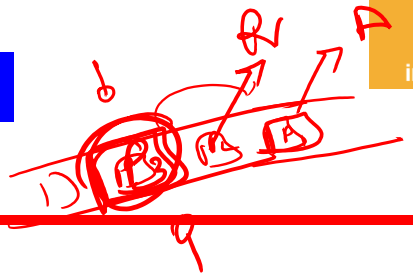
Partially Observable: Poker

Deterministic vs Stochastic

- Deterministic (vs. stochastic): The next state of the environment is completely determined by the current state and the action executed by the agent. (If the environment is deterministic except for the actions of other agents, then the environment is strategic)
- Stochastic: uncertainty as to next state. Ex: robot hitting a goal in football, poker, autonomous driving, agricultural robot and rains.



Episodic vs Sequential



- Episodic: The agent's experience is divided into atomic "episodes" (each episode consists of the agent perceiving and then performing a single action), and the choice of action in each episode depends only on the episode itself.
- Example of Episodic:
  - Given an image, find if there is a cat in it.
  - Expert advice systems – an episode is a single question and answer.
- Sequential: if current decisions affect future decisions, or rely on previous ones. Long-term planning is needed.
- Example of Sequential is packman, investment portfolio management bot.

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

Static vs Dynamic

- Static: The environment is unchanged while an agent is deliberating/thinking. Ex: Puzzle, rubik's cube etc.
- Dynamic: Environment can change while agent is thinking. Ex: raining, signals etc.
- Semidynamic: Environment does not change with time, but performance scored does. Ex: chess with timing…

Rubik's Cube

VS.

Football playing robots

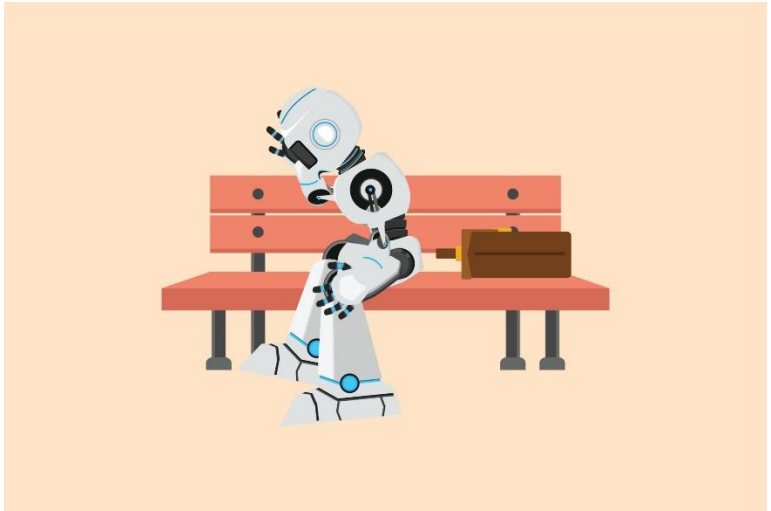
Discrete vs Continuous

- Discrete : A limited number of distinct, clearly defined percepts and actions. Ex: chess game
- Continuous environments have a certain level of mismatch with computer systems. Ex: taxi driving.

VS.

Single Agent vs Multiagent

- Single agent: An agent operating by itself in an environment.
- Multiagent: Presence of other agents. Other agents could be anything that changes from step to step, and can sense and act.



Car-manufacturing robots

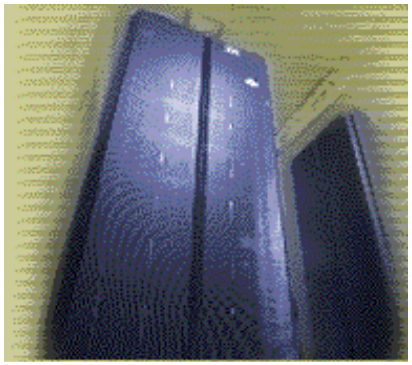
Examples

- The environment type largely determines the agent design
- The real world is (of course) partially observable, stochastic, sequential, dynamic, continuous, multi-agent.

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

BITS Pilani, Deemed to be University under Section 3 of UGC Act, 1956

Activity [5min]
Identify the Environment



Deep Blue

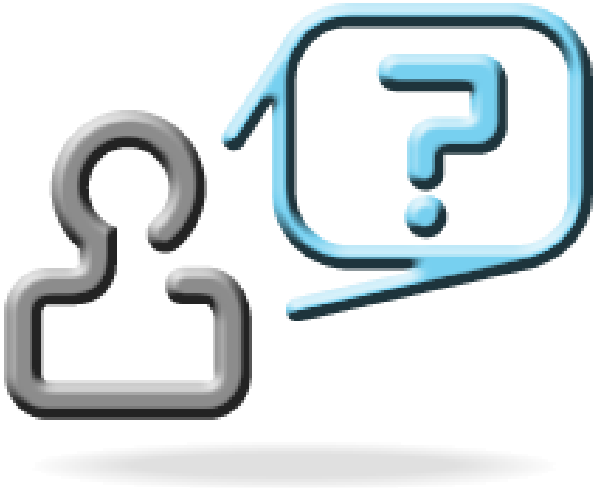
- Static/Semi-dynamic
- Deterministic
- Observable
- Discrete
- Sequential
- Multi-Agent

Robot

- Dynamic
- Stochastic
- Partially observable
- Continuous
- Sequential
- Multi-Agent

Thank You for your time & attention !

Slides are Licensed Under : CC BY-NC-SA 4.0



We will explore more on Rational Agents in the next class …